

Epitope-Based Immunome-Derived Vaccines: A Strategy for Improved Design and Safety

Anne S. De Groot, Leonard Moise, Julie A. McMurry, and William Martin

Abstract Vaccine science has extended beyond genomics to proteomics and has come to also encompass ‘immunomics,’ the study of the universe of pathogen-derived or neoplasm-derived peptides that interface with B and T cells of the host immune system. It has been theorized that effective vaccines can be developed using the minimum essential subset of T cell and B cell epitopes that comprise the ‘immunome.’ Researchers are therefore using bioinformatics sequence analysis tools, epitope-mapping tools, microarrays, and high-throughput immunology assays to discover the minimal essential components of the immunome. When these minimal components, or epitopes, are packaged with adjuvants in an appropriate delivery vehicle, the complete package comprises an epitope-based immunome-derived vaccine. Such vaccines may have a significant advantage over conventional vaccines, as the careful selection of the components may diminish undesired side effects such as have been observed with whole pathogen and protein subunit vaccines. This chapter will review the pre-clinical and anticipated clinical development of computer-driven vaccine design and the validation of epitope-based immunome-derived vaccines in animal models; it will also include an overview of heterologous immunity and other emerging issues that will need to be addressed by vaccines of all types in the future.

Keywords Epitope · Immunome · Vaccines

1 Introduction

The availability of immunome-mining tools has fueled the design and development of vaccines by a process that has come to be termed ‘reverse vaccinology,’ ‘vaccinomics,’ ‘immunome-derived vaccine’ (IDV) design, or ‘genome-derived

A.S. De Groot

Director, Immunology and Informatics Institute, Center for Vaccine Research and Design, University of Rhode Island; Brown University; CEO, EpiVax, Inc., Providence RI 02903, USA

e-mail: AnnieD@brown.edu

vaccine' design (Rappuoli and Covacci 2003; Petrovsky and Brusic 2002; Pederson 1999; De Groot and Martin 2003; Doytchinova, Taylor and Flower 2003). This vaccine concept is based on the identification of a minimal set of antigens that induce a competent immune response to a pathogen or neoplasm. Recognition of antigens occurs through the presentation of B cell and T cell epitopes derived from the antigen, in the correct immunological milieu. In its minimal form, an IDV would contain only adjuvanted B cell and T cell epitopes in delivery vehicles such as liposomes. When these minimal components are packaged in an appropriate delivery vehicle, the complete package comprises an IDV.

Compared to traditional vaccines, IDVs have the potential to be safer and more effective since the vaccine focuses the protective immune response on the most essential antigenic elements of the pathogen/neoplasm. A number of IDVs have been tested in clinical trials (Elliott 2008; Gahery et al. 2006; Asjö et al. 2002; Kran et al. 2004). Because epitope-based IDVs are generally considered to be safe, when compared to other vectored or attenuated live vaccines, many have progressed rapidly from pre-clinical concept into clinical trials. In the cancer vaccine field, where epitope-based vaccines are well-established, many such vaccines are currently in Phase I/II clinical trials (Pietersz, Pouniotis and Apostolopoulos 2006).

This chapter will review the development and validation of IDVs; it will also provide a step-by-step guide to develop IDV using validated immunoinformatics tools. These tools have the potential for dramatically accelerating the development of new and improved vaccines for the emerging and existing infectious diseases. Whole-antigen-based IDVs will be covered briefly; however, the main focus will be on epitope-based IDVs and a description of the immunoinformatics tools that have been developed to accelerate the pre-clinical phase of vaccine discovery. So as to illustrate the process of pre-clinical vaccine development using these tools, two epitope-based IDVs case studies will be presented: (i) a genome-derived vaccine for Tularemia and (ii) an epitope-based HPV vaccine for adjunctive treatment of cancer.

2 Defining the Immunome

Application of molecular biology techniques led to the sequencing of genomes and improved definition of the proteome (expressed proteins). Even though the fundamental concept of the 'immunome' (the subset of fragments of expressed proteins that interface with the host immune system) is relatively well accepted, many important questions remain. Due to genetic variation and poorly understood determinants of antigen processing, it has become clear that immunomes can vary substantially from one host to the next, even when major histocompatibility complex (MHC) and antibody germline genes are shared. The extent of the overlap between immunomes (in different hosts) and the general size of the immunome-representing epitopes derived from a particular pathogen or cancer

remain to be determined. Recently published studies are beginning to address these questions, partly due to the availability of algorithms that facilitate the identification of epitopes from whole genomes.

2.1 How Large Is the Immunome?

The size of the immunome has been puzzling vaccinologists for decades. Certainly, there are examples in the literature that ‘a single epitope protects.’ For example, Crowe et al. (2006) recently found that immunization with a single Th epitope provided a one-log reduction in influenza viral titers early in infection. A single epitope has also been shown to protect against viral disease in woodchucks (Menne et al. 1997), and multiple single epitopes have protected mice against an array of pathogens (An and Whitton 1997 and Olsen et al. 2000). A more diverse set of T cell epitopes appear to be critical to immune response to vaccinia (in mice): 49 Class I MHC epitopes were shown to contribute to the large majority of the CD8+ T cell immune response (Moutaftsi et al. 2006). Remarkably, it was found that 49 of these predicted epitopes (derived from more than 175,000 candidates) represented over 90% of the vaccinia-specific CD8 T cell repertoire. How this number might be extrapolated to humans is unknown, but the result is relevant because 49 epitopes is few enough to be easily packaged and delivered in a vaccine. An epitope-based IDV for genetically diverse populations of humans will almost certainly require more than that number of epitopes, particularly if the vaccine is intended to protect against complex bacteria or viruses, or against solid tumors presenting variable antigenic profiles. Harnessing the power of immunoinformatics accelerates the tasks of defining the immunome and of identifying and developing new vaccines for human diseases.

3 Steps in the Development of an Epitope-Based IDV

The task of developing an epitope-based IDV can be deconstructed into a series of achievable steps.

3.1 Select Protein Antigens of Interest

After selecting a target organism, the next step in the immunome-to-vaccine process is to identify, from within the target genome, a set of potentially antigenic genes/proteins. These sequences can then be screened using a variety of *in silico*, *in vitro*, and *in vivo* mechanisms or a combination of methods. In the case of small viral genomes, it may be possible to include the entire genome in the search universe. In the case of larger bacterial and viral pathogens, the traditional

vaccine targets include surface proteins and secreted proteins (which can easily be found using *in silico* screening programs), toxins (which can be identified through homology matching), and virulence factors (which can be identified through the use of comparative genomics). However, many other types of proteins may also be worthy of consideration: in particular, proteins expressed in high amounts (such as viral capsid proteins), proteins overexpressed during growth or replication, or proteins overexpressed in response to stress conditions.

For cancer vaccines, comparisons between cancerous and normal tissue can uncover cancer-specific genes. One approach involves the use of mRNA derived from cancerous and normal tissue to probe DNA microarrays, followed by selection of genes that are upregulated in cancerous and not in normal tissue (Mathiassen et al. 2001; Sepkowitz 2001).

Using 'Reverse Vaccinology', a term recently coined by Rino Rappuoli, the immune memory of subjects who have successfully encountered and defeated a pathogen can be interrogated to identify the primary targets of a natural immune response. Rappuoli and colleagues identified novel vaccine targets using *in silico* techniques to screen the target genome for 'surface protein-like' sequences. Candidate proteins were then expressed in *Escherichia coli* and screened against human sera isolated from pathogen-exposed subjects. Reactive proteins were deemed relevant to immune response (Pizza et al. 2000).

A related approach for discovering candidate vaccine antigens involves analyzing the target pathogen's proteome *in silico*, using T cell epitope mapping tools. Putative T cell epitopes identified can then be screened against peripheral blood mononuclear cells (PBMC) isolated from human subjects who have been infected with the target pathogen (or who have the target cancer). T cell reaction to a particular peptide epitope, typically measured by ELISA or ELISpot assay, implies that the protein from which the peptide was derived, expressed, processed, and presented to the immune system in the course of a 'natural' immune response. Using this method, measuring immune response to an epitope reveals a protein antigen. Our group describes this approach as 'fishing for antigens using epitopes as bait.'

Other common *in vitro* techniques used to identify expressed or over-expressed genes/proteins include: 2D SDS-PAGE, (Kaufmann et al. 1992; Sonnenberg and Belisle 1997; Hernychova et al. 2001), mass spectrometry (Tomlinson, Jameson and Naylor 1996), and/or tandem mass spectrometry (Barnea et al. 2002).

In the context of vaccines against infectious diseases, it may be prudent to exclude proteins that are highly conserved across species; such proteins (including housekeeping genes) may cross-react with unrelated avirulent organisms or with self-proteins to which there may be pre-existing tolerance. However, conservation *across species* should not be confused with conservation *within species variants*. Sequence conservation within species variants is a highly desirable trait for vaccine components and one that the IDV approach is particularly well suited to harness. RNA viruses in particular (HCV, HIV, coronaviruses) are highly variable pathogens. In these cases, selecting epitopes

that are conserved across variants or subtypes may allow for the development of a more broadly applicable vaccine. Alternatively, single proteins that are relatively well conserved, when compared to the balance of the target pathogen, can be selected as vaccine candidates. Recently, for example, one team has developed a hexon-epitope vaccine that may be effective against a range of adenoviruses, across serotypes (Leen et al. 2008).

Once critical antigens have been identified, the next steps in epitope-based IDV development are to select epitopes and confirm their immunogenicity.

3.2 Identifying B Cell Antigens

Once the adaptive immune system has been engaged, a humoral, or antibody-based response forms the first line of defense against most viral and bacterial pathogens. Antibodies recognize B cell epitopes composed of either linear peptide sequences or conformational determinants, which are present only in the three-dimensional form of the antigen. Several B cell epitope prediction tools, such as 3DEX and CEP, have been proposed and are in the process of being refined (Enshell-Seijffers et al. 2003; Schreiber et al. 2005; Kulkarni et al. 2005). Unfortunately, the computational resources and modeling complexity required to predict B cell epitopes are enormous. This complexity is due in part to the inherent flexibility in the complementarity determining regions (CDR) of the antibody and in part due to glycosylation, and other post-translational modifications can result in modification of B cell epitopes.

Although accurate B cell epitope mapping tools remain elusive, the selection of potent B cell antigens can be accelerated using T cell epitope mapping tools. When considering B cell antigens as potential subunit vaccines, it may be important to also consider their T cell epitope content since the quality and kinetics of the antibody response is dependent upon the presence of T help. B cell antigens, which contain a significant T help, may outperform B cell antigens lacking cognate help. And in some cases, an identified T cell epitope may contain a B cell epitope. Although different epitopes activate T and B cells, it has been widely reported that B cell epitopes have been shown to co-localize near, or overlap, Class II (Th, CD4+) epitopes (Graham et al. 1989; Rajnavolgyi et al. 1999).

3.3 Identifying T Cell Antigens

The adaptive immune system's second line of defense is the T lymphocyte. Class I-restricted cytotoxic T cells (CD8+ CTL) directly engage and attack infected host cells. Class II-restricted T helper cells mediate the growth and differentiation of both T effector cells and antibody-producing B lymphocytes. Both Class I and Class II T cells carry out their roles in response to T cell epitopes, small linear fragments derived from protein antigens, displayed on the surface of

antigen-presenting cells (APC) by various alleles of MHC. While B cells and antibodies generally recognize epitopes on surface proteins only, T cells recognize epitopes derived from a variety of proteins.

Once taken up by APC, antigenic proteins are broken down by digestive enzymes. During this process very large numbers of peptide fragments are released. Any one of these fragments could be a T cell epitope, but only about 2% of all the fragments generated can implant themselves in the binding groove of the MHC molecule and be presented on the surface of the APC. One of the critical determinants of T cell epitope immunogenicity is the strength of epitope binding to MHC molecules (Lazarski et al. 2005). Peptides binding with higher affinity are more likely to be selected by MHC molecules and to be displayed on the cell surface where they can be recognized by T lymphocytes. Using a variety of methods including frequency analysis, support vector machines, hidden Markov models, and neural networks, researchers have developed highly accurate tools for modeling the MHC–peptide interface and for accurately predicting T cell epitopes. For a review of T cell epitope mapping tools, see De Groot and Berzofsky (2004) and the accompanying issue of *Methods*. What all these tools have in common is an ability to quickly screen large volumes of genomic sequences for putative epitopes; this preliminary screen reduces the search space dramatically, typically by at least 20-fold.

The ability to accurately predict T cell epitopes from raw genomic data is fundamental to the development of an IDV. However, even a highly accurate prediction is still only a prediction. Before including predicted epitopes in a candidate vaccine, it is important to validate their immunogenicity *in vitro* and *in vivo*.

3.3.1 In Vitro Assays: Peptide Binding Assays

Once identified, peptides representing the selected epitopes are then synthesized. HLA binding assays can be used to assess whether peptides derived from immunoinformatics analysis can bind to either MHC Class I or Class II by measuring the affinities of predicted epitope sequences for the HLA alleles *in vitro*. *In vitro* evaluation of MHC binding can be performed by quantifying the ability of exogenously added peptides to compete with a fluorescently labeled known MHC ligand (Steere et al. 2006) and can be adapted for high throughput (McMurry et al. 2007a). EpiVax routinely uses these high-throughput HLA binding assays to confirm epitope predictions *in vitro*. A concordance between HLA binding and immunogenicity is often observed (McMurry et al. 2005).

3.3.2 In Vitro Assays: Measuring T Cell Responses

Peptides are used to measure T cell responses *in vitro*; they can be of variable lengths (9–25). Peptides presented in the context of Class I MHC are generally limited to 9 or 10 amino acids in length, although some processing is believed to occur during the T cell assay and so 15-mers are also used for Class I assays.

In contrast with Class I epitopes, which are short and fit tightly in the bounded MHC molecule, Class II (T helper) epitopes lie within an open-ended groove in the MHC II. As such, a Class II epitope can shift within the groove, thereby accommodating MHC of various haplotypes. The only limit on the size of the peptide is its ability to remain in a linear conformation in the open-ended groove.

Both MHC Class I- and MHC Class II-restricted epitopes (targeting CD4+ and CD8+ T cells, respectively) are believed to be important for the development of effective vaccines. CD4+ T helper cells enhance and amplify cytotoxic T cell (CTL) immune responses and have been shown to be important in the development of CD8+ T cell memory to a range of pathogens (Ahlers et al. 2001). CTLs generally play a role in the containment of viral and bacterial infection (Plotnicky et al. 2003), and the prevalence of CTLs usually correlates with the rate of pathogen clearance.

Like peptides, whole antigens too can be used to measure T cell responses in vitro. The recognition of these antigens requires the presence of an APC that is capable of processing and presenting peptides derived from the antigen.

If blood from exposed individuals is available, the peptides validated as MHC ligands in binding assays can be tested for their reactivity with T cells, serum, or both. A positive immune response (as measured by ELISA, ELISpot, or intracellular cytokine staining) should be interpreted as a sign that the parent protein interfaces with host immune response in the course of natural infection or disease. Following confirmation, the peptides that stimulate a response can be considered vaccine candidates themselves or can be used to select the entire protein for use in a subunit vaccine. These candidates can then be incorporated into a vaccine delivery vehicle with an appropriate adjuvant.

ELISA and ELISpot are related methods for detecting T cell responses by the measurement of cytokines secreted by the T cells (gamma interferon, IL-2, and IL-4 are examples). The expansion (proliferation) of T cells in response to stimulation by peptide:MHC can be measured by (1) the dilution of a fluorescent dye in subsequent generations of cells (CFSE) and (2) the incorporation of a radioactive label in the proliferating cell's DNA (tritiated thymidine incorporation assay). Fluorescence activated cell sorting (FACS) and intracellular cytokine staining (ICS) are the most precise methodologies available for measuring and defining T cell response. For example, T cells that respond to a particular epitope can be directly labeled using tetramers (comprising MHC Class II: peptide complexes). Labeled cells can then be sorted and counted, and the phenotype of T cells that respond to the antigen can be determined using cell surface markers and ICS (Tobery et al. 2006).

3.4 Select Delivery Vehicle and Adjuvant

Factors extrinsic to processing, such as the cytokine milieu induced in response to a particular component of a vaccine (Krieg et al. 1998) or pathogen

(Ghosh et al. 1998), also play a role in the conditioning of the immune response. Thus, T cell epitopes may be necessary to drive immune response, but are not sufficient. Co-stimulatory molecules that provide a second signal, the right cytokine milieu and other factors directing the nature (Th1 vs. Th2) of the immune response, are also crucial (Shahinian et al. 1993; Kuchroo et al. 1995). Adjuvants provide this added 'boost' in the context of vaccines.

The same range of delivery vehicles that exist for conventional vaccines can be used for the development of IDVs and epitope-based IDVs. For example, IDVs and epitope-based IDVs can be formulated and delivered as pseudo-proteins or peptides in a carrier vehicle such as a liposome or viral-like protein (VLP); alternatively, the sequence of the IDV antigens or epitope string can be inserted into a viral or bacterial vector such as adenovirus or salmonella; alternatively, a DNA vaccine construct encoding the antigen(s) or epitopes can be developed. The choice of adjuvants for use in humans is relatively extensive and each adjuvant has advantages and disadvantages. The advantages and disadvantages of each type of vaccine delivery vehicle and adjuvant here is beyond the scope of this chapter; readers are referred to a review by Fraser et al. (2007).

3.5 Animal Model for Vaccine Efficacy

The next step in the development of epitope-driven IDV is to determine whether immunization provides competent immune response. The IDV or epitope-based IDV is administered and immune responses to the components are evaluated following immunization. Even though a range of animal models are used for the evaluation of vaccines, results from immunogenicity studies in these models should be interpreted with caution. Although their functions may be similar, the MHC of mice, rodents, and non-human primates differ from human MHC (known as HLA in the context of human immune response) at the amino acid level and these differences effect which epitopes can be presented. This helps to explain why different strains of mice (Balb/C, C57Bl/6) have different immune responses to pathogens as well as vaccines for those pathogens (Klitgaard et al. 2006). In particular, epitope-based vaccines that are developed using predicted human T cell epitope mapping tools can be tested only in murine models that are HLA transgenic.

Fortunately, a number of transgenic mouse strains that express the most common HLA A, HLA B and HLA DR, molecules have been developed. T cell responses in these mice correlate directly with T cell responses observed in infected/vaccinated humans (Man et al. 1995; Shirai et al. 1995). HLA transgenic mice are now routinely used to assay and optimize (human) epitope-driven vaccines in pre-clinical studies (Ishioka et al. 1999; Charo et al. 2001; Livingston et al. 2001). Despite the limited number of HLA Class II alleles for which Tg mice have been developed, comparisons of immunogenicity can be done to a high degree of accuracy in the mouse model for selected HLA Class II

alleles (HLA DR 0101, 0301, 0401, 1501). Unfortunately, it appears these mice may have difficulty breeding due to poorly understood consequences of their transgenic heritage, limiting the use of this important model system.

3.6 Challenge Studies

The final step in the development of any vaccine is experimental validation of the immunogenicity and protective efficacy of computationally selected antigens. Currently, a series of experimental vaccines have shown efficacy in animal models and several IDVs are being tested in clinical studies. In the context of infectious disease, genome-derived vaccines that have progressed furthest along the vaccine development pipeline are generally based on whole proteins rather than epitopes (Rappuoli and Covacci 2003; Pizza et al. 2000). However, epitope-based IDVs are currently being developed for a range of infectious diseases by the authors' laboratory and by many others (Depla et al. 2008).

While it is common knowledge that subunit-based vaccines can protect against infection, similar success with epitope-based approaches is not as widely known. In addition to the studies previously cited, immunization of BALB/c mice with three doses of a peptide construct containing an H-2(d)-restricted cytotoxic T lymphocyte (CTL) epitope from a murine malaria parasite induced both T cell proliferation and a peptide-specific CTL response mediating nitric-oxide-dependent elimination of malaria-infected hepatocytes *in vitro*, as well as partial protection of BALB/c mice against sporozoite challenge (Franke et al. 2000). In a separate study, immunization of BALB/c and CBA mice with measles virus CTL epitopes resulted in the induction of epitope-specific CTL responses and conferred some protection against encephalitis following intracerebral challenge with a lethal dose of virus (Schadeck et al. 1999). These are just a few successful examples of many studies carried out in animal models; however, translation to prevention of disease in humans has been difficult to achieve.

3.7 Clinical Development

In contrast with whole-protein subunit vaccines, IDVs and epitope-based IDVs have taken longer to make the transition from animal model to the clinic, mainly due to the novelty of the concept and perhaps unfounded concerns that epitopes are not sufficient for the generation of effective immune response. Cancer therapy is an exception to this rule. As previously described, epitope-based IDVs have been evaluated in the context of therapy against chronic infection or cancer (Ueda et al. 2004; Valmori et al. 2003).

In the cancer vaccine field, where the concept of epitope-driven vaccines is well established, many more peptide vaccines have successfully passed pre-clinical

tests and are currently in Phase I/II clinical trials (Pietersz, Pouniotis and Apostolopoulos 2006). New approaches are emerging, which may improve the success rate and, indeed, the results from recent clinical trials prove the principle. One approach is to identify epitopes that are unique to the tumor (prostate, lung, colon) and to pre-screen the patient for response to the peptide. This approach, called personalized vaccination, takes into account the diversity of CTL epitope recognition among patients. Whereas the response rates to classical (non-personalized) peptide vaccines have been disappointing, responses to personalized vaccines (in a Phase I trial, conducted in Japan) have been as high as 11.1% in the advanced cancers and equal to or more than 20% in malignant glioma and cervical cancers, respectively (Itoh and Yamada 2006).

It is noteworthy that just a few epitope-driven vaccines against viral and microbial pathogens have reached the stage of Phase I or II efficacy trials in humans. For example, Bionor Immuno's HIV p24 gag peptide vaccine (Vacc-4X) was demonstrated to be safe and well tolerated in Phase I trials (Asjö et al. 2002) and dose-dependent and immunogenic in Phase II trials in Norway (Kran et al. 2004). Similarly, Nardin's epitope-based vaccine for malaria is moving along the clinical trial pathway (Nardin et al. 2000).

4 Epitope Mapping Tools for IDV

In this section, we describe the immunomics tools developed and used by the EpiVax vaccine development group in recent collaborations with Dr. Steve Gregory of Lifespan, Dr. Ousmane Koita of the University of Bamako, Mali and Dr. David Weiner of University of Pennsylvania.

4.1 EpiMatrix: T cell Epitope Mapping for IDV

T cell epitopes are linear peptides that bind to MHC molecules. Binding is mediated by the interactions between the R-groups of the amino acids in the peptide ligand and the pockets on the floor of the MHC binding groove. Because the MHC:peptide interaction is well characterized, pattern-matching algorithms can be used to screen protein sequences for peptides that will bind MHC. The authors of this report currently use the EpiMatrix system, a suite of epitope-mapping tools that has been validated by more than a decade of use in selecting putative epitopes for in vitro and in vivo studies (see references (De Groot et al. 1997; Bond et al. 2001; Dong et al. 2004; McMurphy et al. 2005; Koita et al. 2006)). The EpiMatrix algorithm is based on a set of Class I and Class II HLA matrices wherein individual frequencies of all 20 amino acids (aa) in each HLA pocket position are applied to the prediction of overlapping 9- and 10-mer peptides. In a typical analysis, protein antigens are parsed into

overlapping 9-mer frames where each 9-mer overlaps the last by eight amino acids. Each 9-mer is then scored for predicted binding affinity to one or more Class I or Class II HLA alleles. In order to compare potential epitopes across multiple HLA alleles, EpiMatrix raw scores are converted to a normalized 'Z' scale. Peptides scoring above 1.64 on the EpiMatrix 'Z' scale (typically the top 5% of any given sample) are likely to be MHC ligands. Since Class II epitopes can be promiscuous, our approach to the prediction of Class II epitopes is to estimate the binding potential of each frame with respect to each of a panel of eight common Class II alleles (DRB1*0101, *0301, *0401, *0701, *0801, *1101, *1301, and *1501). Taken together, these alleles 'cover' the genetic backgrounds of most humans worldwide (Southwood et al. 1998) and they also represent the predominant types of 'pockets' for the most common MHC.

Recently, the EpiMatrix system has been utilized to measure the potential immunogenicity of whole proteins. In this context, EpiMatrix assesses the aggregate epitope density of a given protein with respect to the aggregate epitope density of a set of randomly generated pseudo-protein sequences of similar size (De Groot 2006). By correcting for the size and expected epitope density, the potential immunogenicity candidate vaccine antigens can be directly compared. Further immunogenicity of low scoring proteins may be enhanced by modifying the immunogenic region amino acid sequence so that it contains more T cell epitopes (see illustration of this approach in the HPV vaccine section, below).

4.2 ClustiMer: Finding Promiscuous T cell Epitopes

Peptides predicted to bind to multiple HLA alleles are known as promiscuous T cell epitopes. The ClustiMer algorithm is used to scan the output produced by the EpiMatrix and identifies the polypeptides predicted to bind to an unusually large number of HLA alleles. Briefly, the scores of each analyzed 9-mer are aggregated. High-scoring 9-mers are then extended at the N- and C-terminal flanks until the predicted epitope density of the promiscuous epitope falls below a given threshold value. This particular approach to mapping epitopes has also been useful for discovering 'EpiBars,' which may be a signature feature of highly immunogenic, promiscuous Class II epitopes. An example of a promiscuous T cell cluster containing an EpiBar (Tetanus Toxin 830–844) is shown in Fig. 1.

A single T cell epitope 'cluster' usually ranges from 9 to about 25 amino acids in length and can contain anywhere from 4 to 40 binding motifs. Using EpiMatrix as described above and ClustiMer, scores above 10 and, in particular, scores above 15 indicate significant immunogenic potential (De Groot 2006). Note the horizontal bar of high Z scores at position 308 in Fig. 3. Having observed this 'EpiBar' pattern to be characteristic of promiscuous epitopes, the authors have integrated the pattern into the prospective selection of clusters.

Tetanus Toxin Peptide AA 830-844

Frame Start	AA Sequence	Frame Stop	DRB1*0101 Z score	DRB1*0301 Z score	DRB1*0401 Z score	DRB1*0701 Z score	DRB1*0801 Z score	DRB1*1101 Z score	DRB1*1301 Z score	DRB1*1501 Z score	HITS	Cluster Score
830	QYIKANSKF	838										
831	YIKANSKFI	839	2.68		1.81	3.03				1.64	4	} Deviation from expectation: 23.82
832	IKANSKFIG	840	1.38	2.83			2.79	2.12	3.13	2.18	5	
833	KANSKFIGI	841										
834	ANSKFIGIT	842										
835	NSKFIGITTE	843										
836	SKFIGITEL	844				1.41						

Fig. 1 EpiBar-Typical EpiMatrix analysis. Z score (Top 10% Top 5% Top 1%) indicates the potential of a 9-mer frame to bind to a given HLA allele. All Z scores in the Top 5% (>1.64) are considered 'hits'. Though not hits, scores in the top 10% are considered elevated; scores below 10% are masked for simplicity. Frames containing four or more alleles scoring above 1.64 are colloquially referred to as 'EpiBars' (see frames 831:YIKANSKFI and 832:IKANSKFIG). This band-like pattern is characteristic of promiscuous epitopes. The tetanus toxin peptide scores are extremely high for all eight alleles in EpiMatrix; the deviation compared to expectation is + 23.82

Promiscuous epitopes also exist, to a certain degree, for Class I alleles. Some laboratories have demonstrated cross-presentation of peptides within HLA 'superfamilies' (such as the A3 superfamily: A11, A3, A31, A33, and A68) described by (Sette and Sidney 1998). The authors have confirmed cross-MHC binding and presentation to T cells in our HIV vaccine studies (De Groot et al. 2001).

5 Additional Vaccine Design Tools

5.1 Conservatrix: Finding Conserved T Cell Epitopes

One limitation of conventional vaccination, and to a lesser extent natural infection, is that the immune system focuses strongly on the most mutable immunogen of the virus – typically, the viral envelope. In the case of HIV and other viruses, vaccination with more conserved, subdominant epitopes has been shown to circumvent this hierarchy and potentiate cross-strain protection (Ostrowski et al. 2002; Nara and Lin 2005). In like manner, a conserved T helper-directed vaccine may provide a more 'democratic' way of stimulating immune response, increasing the number targets for T cell recognition, thereby providing T help to antibody response despite potential viral variability (Santra et al. 2002; Subbramanian et al. 2003; Scherle and Gerhard 1988; Scherle and Gerhard 1986; Russell and Liew 1979; Johansson et al 1987).

The genetic variability of some pathogens constitutes a significant challenge to the efforts to design a vaccine driven by cellular immune response (De Groot et al. 2001; De Groot et al. 2002). The authors have been involved in developing an HIV-1 vaccine that includes highly conserved (cross-clade) T cell epitopes. The Conservatrix algorithm, developed for this application, parses input sequences into component strings (the lengths of the strings may be determined

by the operator) and then searches the input dataset for matching segments. Conservatrix may be used to compare strings derived from different strains of the same organism (hepatitis C, for example, or HIV) or to search a given sequence for a user-supplied target sequence. Target sequences may be input as specific sequences or as coded patterns. Thus, the operator can use 'wild cards', allowing for one or more of the amino acid residues in any given peptide sequence to be any amino acid [X], or a limited set of amino acids such as [L, V]. Results of each analysis are stored in a database and may be browsed or exported to another program for analysis.

By selecting highly conserved epitopes, regardless of their distance from the ancestral HIV-1 genome, we have identified sequences conserved for structural and functional reasons and are therefore less likely to be modified in the course of further evolution of HIV-1 (Peyerl et al. 2004; Koibuchi et al. 2005).

5.2 EpiAssembler: Immunogenic Consensus Sequence Epitopes

The problem of virus variability also significantly complicates the selection of epitopes that have a population-coverage advantage; such epitopes are termed 'clustered,' 'superfamily,' or 'promiscuous.' To address this problem, the authors developed EpiAssembler (De Groot et al. 2004) to identify sets of overlapping, conserved, and promiscuously immunogenic epitopes and assemble them into extended immunogenic consensus sequences (ICS) (see Fig. 2).

In theory, proper processing and presentation of these sequences would allow for the presentation of highly conserved peptides in the context of more than one MHC. The resulting peptide is not a 'pseudosequence' as such, since each constituent epitope occurs in its corresponding position in the native protein. Thus, while the full-length 'immunogenic consensus sequence' is not necessarily found in any one variant sequence, the peptide is more representative of the sequence universe.

In the case of HIV, for example, we used the ICS approach to design a peptide-based vaccine. While the full-composite ICS peptides happen to be exactly conserved in a few individual strains of HIV, each peptide represents a significant percentage of circulating strains, because every constituent overlapping epitope is conserved in a large number (range 893 to 2,254) of individual HIV-1 strains. As compared with immunogenic consensus sequences, randomly selected counterparts, on average, contain half as many binding motifs and cover a third fewer isolates. To develop vaccines of equivalent antigenic 'payload,' using conventional methods would be prohibitively expensive as it would require including multiple different variants of each antigen. This approach has been useful for identifying highly immunogenic epitopes for HIV vaccine design (De Groot et al. 2005). By focusing on conserved, MHC-promiscuous T helper epitopes, the ICS approach has the potential to efficiently overcome the genetic variability of both virus and host.

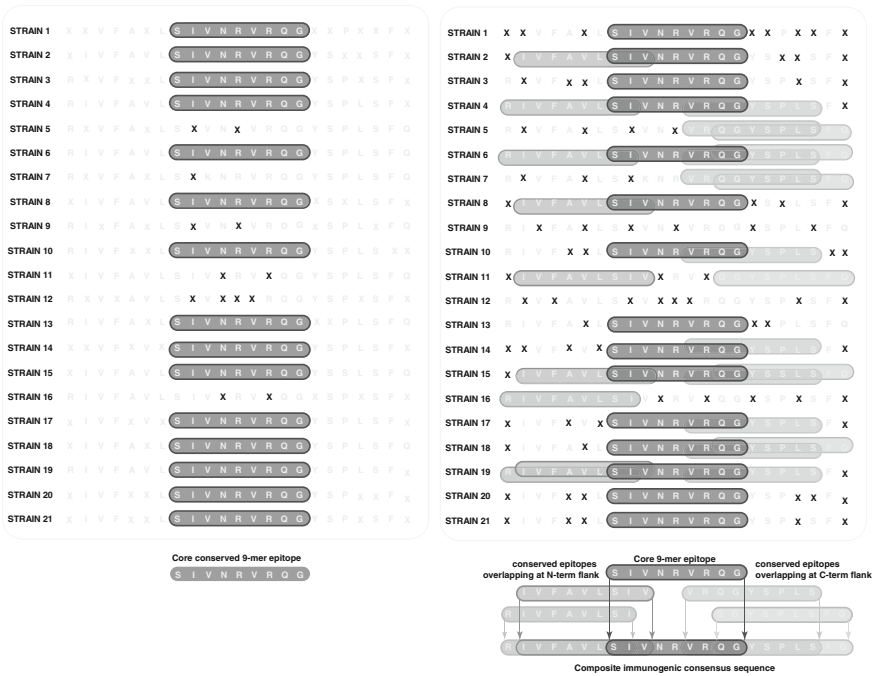


Fig. 2 The ICS assembly operation was performed using EpiAssembler (Bill Martin, EpiVax, 2004). *Left panel:* each variant strain is first analyzed and a highly conserved, putatively promiscuous 9-mer is chosen as the core peptide. Mismatches with the selected epitope sequences are represented with the letter X. *Right panel:* Additional epitopes are then identified, which overlap with the natural N- and C-terminal flanking regions of the core 9-mer epitope. The overlap length requirement is decreased by one amino acid iteratively until reaching a minimum of three overlapping amino acids. If more than one suitable overlap is identified, the overlapping peptide with the higher overall EpiMatrix rank is selected. This process is repeated using the extended peptide as the new core sequence. The cycle can be repeated for the length of an entire protein or can be truncated when the peptide reaches a length that can be easily produced synthetically

5.3 Eliminating Cross-Reactivity (BlastiMer)

One of the advantages of IDV is that it is possible to omit deleteriously cross-reactive epitopes. Perhaps, the most famous example of an adverse effect due to cross-reactivity with self was observed following vaccination for Lyme disease with the Osp A protein. The vaccine has been recently re-engineered with the cross-reactive epitope removed (Willett et al. 2004). In the context of our own work, peptides selected for in vitro evaluation are evaluated for homology with human proteins by BLASTing the sequences against the human sequence database at GenBank (<http://www.ncbi.nlm.nih.gov/>).

BLASTiMer automates the process of submitting sequences to the websites featuring search engines such as the blast engine at NCBI (www.ncbi.nlm.nih.gov/blast). By default, BLASTiMer blasts sequences against all non-redundant

GenBank CDS translations, PDB, SwissProt, PIR, and PRF. BlastMer assesses the homology between the submitted sequence and the sequence of proteins of other organisms. Patent BLAST, on the other hand, targets a database of sequences gleaned from patents. Users of either program may control all of the submission options available to interactive users at NCBI. In both cases, results are recorded in a database and can be browsed, exported, or summarized and rendered in a report format. According to the authors' standard practice, any peptide that shares greater than 80% identity with peptides contained in the human proteome is eliminated from consideration in a vaccine.

5.4 Vaccine CAD: Aligning Epitopes

A number of methods for enhancing epitope-based vaccines have been described and implemented (Thomson et al. 1998; Rodriguez and Whitton 2000). One approach is to align the epitopes in a protein or DNA vaccine construct as a 'string of beads' without any intervening sequences or 'spacers' in a DNA plasmid encoding the individual epitopes (An et al. 2000). However, the lack of 'natural flanking sequences' – has raised concern that their proteolytic processing may be compromised, and that junctional epitopes, peptides other than the specific peptides of interest, may be generated as a result of processing (Godkin et al. 2001). To address this concern, the authors developed Vaccine-CAD (see Fig. 3), an algorithm that incorporates the evaluation of

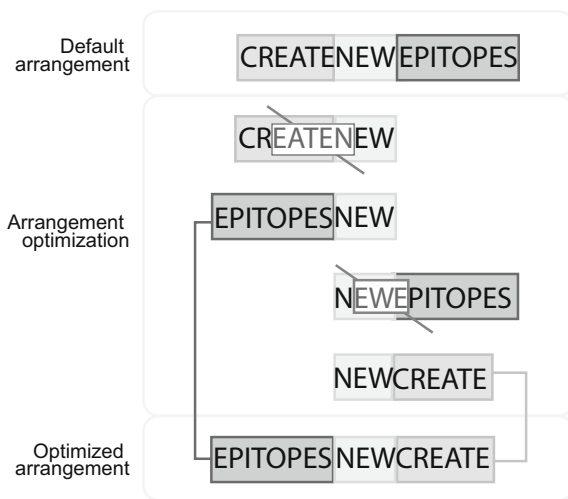


Fig. 3 Vaccine-CAD, illustrated with three sample epitopes represented by the words 'create,' 'new,' and 'epitopes.' The default arrangement of the words results in unintended sequences, represented by the words 'eaten' and 'ewe,' at the junctions between the intended epitopes. Iterative modifications in the arrangement of the epitopes results in the development of a sequence that has no 'pseudoepitopes' (new epitopes that were not intended) at the junctions of the juxtaposed epitopes

junctional epitopes, the insertion of spacers and breakers, the requirements for secretion or processing tags, and the evaluation of epitope strings for potential homologies to human protein fragments.

5.5 HLA Coverage for IDV

An important component in the epitope-driven vaccine process is the selection of epitopes from the regions of pathogens that are presented by MHC molecules for T cell recognition. ‘MHC binding motifs,’ first identified by Rötzschke and Falk, are the patterns of amino acids in peptides that are known to promote the binding of peptides containing these patterns to the MHC molecules on the surface of APCs (Falk et al. 1991; Rotzschke et al. 1991). Different MHC molecules have different binding motifs, limiting the set of MHC ligands that can be presented in the context of any given MHC.

While consideration of HLA alleles may lead to concern about the selection of epitopes for broad coverage of populations, Gulukota and DeLisi (1996) and Sette and Sidney (1998) have demonstrated that epitope-based vaccines that contain epitopes restricted by selected ‘supertype’ HLA can provide the broadest possible coverage of the human population. Furthermore, recent studies by Brander and Walker indicate that there may be even greater flexibility in the binding of epitopes to MHC than previously recognized; this is also consistent with the data recently presented by Frahm et al. (2004). The inclusion of ‘promiscuous epitopes’ – epitopes that are recognized in the context of more than one MHC (Paina-Bordignon et al. 1989; De Groot et al. 1991; Sette and Sidney 1998) in epitope-driven vaccines may therefore overcome the challenge of genetic restriction of immune response. In addition, the repertoire of possible MHC-restricted epitopes recognized by an individual’s T cells has been shown to be quite variable, even between HLA-matched individuals (Jameson, Cruz and Ennis 1998; Gianfrani et al. 2000; Betts et al. 2001).

5.6 Aggregatrix: Aggregation of Epitopes into the Ideal IDV

Aggregatrix, a new algorithm that was recently developed at EpiVax, iteratively searches for the combination of epitopes that achieves maximal cross-clade representation. The authors performed this analysis for our HIV epitopes, as shown in Fig. 4, evaluating each HLA-A2-restricted HIV peptide individually and the set in aggregate for coverage of HIV-1 strains by year (1995–2005), by country of origin, and by clade. As can be seen in this figure, the set of highly conserved A2-restricted peptides we have tested and confirmed in ELISpot assays covered between 54% and 86% of strains in a given year, between 33 and 100% of strains in a given country, and between 0% and 100% of strains in a given clade. The HLA A2 peptides cover 85, 78, 78, and 80% of clades A, B, C, and D, respectively. This is a remarkable breadth of coverage for a limited set of

Coverage of GAIA HIV B7 peptides--individually and in aggregate--
across time, countries, and clades

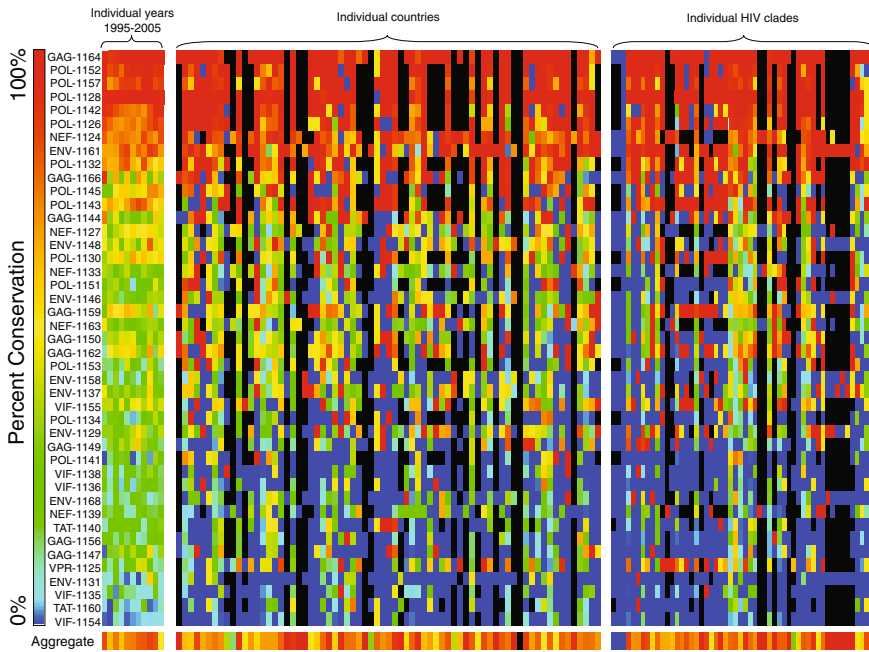


Fig. 4 GAIA HIV A2 peptides — individually and in aggregate — percentage coverage of strains by year, country, and clade. Each row of the matrix denotes a specific peptide; the peptide’s protein of origin is included within the peptide ID. Each column of the matrix denotes a specific year, country, or clade, grouped as indicated. The percentage coverage of strains is represented on a color gradient, with warm tones indicating values above 50% and cool tones indicating values below 50%. The *bottom* row of the matrix shows the percentage of each protein set that contains at least one peptide from our pool. *Black boxes* indicate that no isolates of the protein are available for that year, clade, or country. The *bottom* row represents the aggregate percent coverage for the epitope set. Each cell of the matrix represents the percentage coverage per peptide, except for the bottom row cells, which represent the aggregate percentage coverage for the peptide set. Column headers are listed here for space considerations: left to right, the year columns are 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, and 2005; aggregate coverage of strains by year ranges from 47% (2000) to 63% (2004). The countries left to right are: Angola, Argentina, Australia, Belgium, Benin, Bolivia, Brazil, Botswana, Belarus, Canada, The democratic republic of the Congo (Zaire), Congo, Cote d’Ivoire, Chile, Cameroon, China, Colombia, Cuba, Cyprus, Germany, Djibouti, Dominican Republic, Ecuador, Estonia, Spain, Ethiopia, France, Gabon, United Kingdom, Georgia, Ghana, Gambia, Equatorial Guinea, Greece, Hong Kong, Israel, India, Italy, Japan, Kenya, Republic of Korea, Mali, Myanmar, Namibia, Niger, Nigeria, Netherlands, Norway, Portugal, Russian federation, Rwanda, Sweden, Senegal, Somalia, Chad, Thailand, Trinidad and Tobago, Taiwan, United Republic of Tanzania, Ukraine, Uganda, United States, Uruguay, Uzbekistan, Venezuela, Vietnam, Yemen, South Africa, Zambia, and Zimbabwe; aggregate coverage of strains by country ranges from 33% (Gabon) to 94% (Italy). The clade columns left to right are: 0206, 0209, 1819, A, AAD, AB, AC, ACD, ACDGKU, AD, ADF, ADGU, ADHK, AE, AF, AFG, AFU, AG, AGH, AGHU, AGJ, AGU, AHJU, AU, B, BC, BF, BG, C, CD, CGU, CK, CPX, CU, D, DF, DK, DO, DU, F, G, GHJKU, GK, GGU, GU, H, J, JKU, JU, K, N, O, U. The HLA-A2 peptides together cover 51, 51, 53, and 44% of clades A, B, C, and D, respectively. On average, the chimeric clade sequences were 50% covered by our HLA-A2 restricted epitopes (*See Color Insert*)

HLA A2 epitopes, given the well-known ability of HIV to mutate away from HLA (Nguyen et al. 2004; Iversen et al. 2006).

5.7 Individualized T Cell Epitope Measure (iTEM)

In studies of immune response to therapeutic proteins, the authors have observed that subject-to-subject variation in T cell response closely relates to (1) subject HLA type and (2) the number of protein-derived peptides that match the subjects HLA. To describe this relationship, the authors developed a metric that may be useful in the development of epitope-based vaccines or in the clinical assessment of immune response to vaccines, called the ‘individualized T cell epitope measure’ or iTEM. The iTEM can be calculated for each subject who is responding to a given epitope by summing the EpiMatrix Z-scores for each positive peptide for each HLA allele in a given subject’s haplotype, thus:

$$\begin{aligned} \text{iTEM score} = & [\text{EpiMatrix score of peptide for HLA type 1}] \\ & + [\text{EpiMatrix score of peptide for HLA type 2}] + \dots \end{aligned}$$

The same calculation can be performed for larger peptides and proteins by summing all of the scores for the subject’s HLA type. This calculated score allows for the individualized potential immunogenicity to be predicted, based on the number of putative epitopes contained in a protein that might be presented to their T cells, based on their HLA haplotype. Using this score, it is possible analyze the contribution of haplotype to the corresponding T cell response. In our prospective evaluations, significant correlations were found between the IFN-gamma response to a given antigen and the iTEM scores for individual subjects (for data published in Koren et al. (2007) $r = 0.69$, $p = 0.0134$). In addition, correlations between the iTEM score and patient HLA were also observed for their antibody titers. Further evaluations of this method for predicting individual responses to vaccines and therapeutic proteins are in progress.

5.8 Anticipating Processing and Presentation

Defining the immunome by identifying T cell epitopes and confirming their immunogenicity is but the first step of vaccine development. A number of conditions extrinsic to the MHC-ligand interaction may influence the final composition of the epitope ensemble. For example, whether or not a predicted epitope is confirmed is related to: (1) the quantitative expression of the source protein; (2) the number of different epitopes derived from this protein, which are presented on the surface of the APC (Wherry et al. 1999); (3) the amino acids that flank the epitope (Bergmann et al. 1994; Shastri, Serwold and Gonzalez 1995; Livingston et al. 2001); and (4) proper cleavage and trimming by proteolytic enzymes in the proteasome and processing pathways (Van Kaer et al. 1994; York et al. 1999; Chen et al. 2001; Toes et al. 2001). For example, it is very likely

that end-to-end epitope presentation, such as was used in the design of the OXAVI HIV gag/epitope vaccine, impaired the presentation of the epitopes in immunogenicity studies.

Vaccine-CAD also takes into account the role of flanking residues: Studies conducted in murine models have demonstrated that residues flanking an MHC Class I epitope strongly influence the delivery of the intact epitope to TAP following proteasome degradation (Thomson et al. 1995; Hozhutter, Frommel and Kloetzel 1999; Mo et al. 1999). In addition, Livingston et al. (2001) have tested a standard spacer sequence (-GPGPG-) for vaccine constructs consisting of MHC-II-restricted, Th-cell epitopes; the use of this spacer disrupts junctional epitopes that might compete for degradation or for MHC binding (both G and P are unusual carboxy-terminal anchors for a peptide that binds to Class II MHC). This approach has been used for constructs with up to 20 epitopes, in assays where responses were detected to the majority of epitopes (Livingston et al. 2001).

6 Methods of Confirming IDV

6.1 Two Case Studies

6.1.1 Bacterial (Tularemia)

Francisella tularensis is a zoonotic bacterium. It is endemic to certain communities such as Martha's Vineyard, Massachusetts, USA, where it is known as Rabbit Fever. Tularemia represents a potentially dangerous biological weapon owing to its high degree of infectivity, ease of dissemination, and capacity to cause severe illness. Despite several decades of research, no vaccine for tularemia is licensed for public use. For a review of tularemia vaccines, see McMurry et al., 2007b.

We have been actively developing an epitope-based tularemia vaccine combining computational immunology with in vitro and in vivo validation (McMurry et al 2007a). The starting point of our vaccine was the fully annotated *F. tularensis* subsp. *tularensis* (SCHUS4) genome published in by Larsson et al. (2005). A prototype vaccine containing only Class II epitopes has been tested in challenge studies in HLA DRB1*0101 transgenic mice. For this vaccine, the EpiMatrix algorithm was utilized to identify highly promiscuous T cell epitopes within the tularemia genome. Twenty-five Class II-restricted epitopes were selected, synthesized, and screened in vitro using a recombinant soluble HLA Class II competition-binding assay (described above). Peptides that bound with high affinity were then tested ex vivo in ELISpot assays with blood obtained from *F. tularensis* subsp. *tularensis*-exposed individuals. Search-Light analysis was also performed on supernatants derived from human cell culture stimulated with peptide using a panel of nine cytokines. Forty-two percent of peptides bound to DRB*0101 and are likely to bind to several other alleles (in that they were predicted using the ClustiMer algorithm).

ELISpot assays showed positive IFN-gamma responses to 21/25 individual peptides and to peptide pools in nearly all of the 23 human study. The number of epitopes recognized per subject ranged from 1 to 17 and averaged 4 per subject; not every peptide was tested for every subject.

Peptides that elicit a robust memory response, as evaluated by these various assays, were incorporated into a vaccine construct and tested in challenge studies with HLA transgenic mice as a possible vaccine against tularemia. Immunogenicity studies in HLA transgenic DRB1*0101 mice were performed using the multi-Class II-epitope DNA constructs and/or peptides representing the epitopes, and T cell responses were evaluated in IFN-gamma ELISpots. HLA DR1 transgenic mice were challenged with five times the LD50 of *F. tularensis* LVS; 57% of vaccinated mice survived, all non-vaccinated mice died (manuscript in preparation). This result demonstrates the potential for a genome-derived epitope-based vaccine to protection from a Class A bioterror pathogen.

Importantly, the protection we observed is accounted for by only 10 of the 14 SCHUS4 epitopes in the vaccine, as they were conserved in the LVS challenge strain. The SCHUS4-specific epitopes that were found to be significantly immunogenic would further contribute to protection in a SCHUS4 challenge. This result is consistent with other findings that a limited set of epitopes may be sufficient to induce a protective immune response (Moutaftsi et al. 2006). Studies using SCHUS4 (the wild type tularemia) are planned. The results obtained to date appear to indicate that vaccine design that originates with the whole genome may lead to the development of a protective epitope-based vaccine.

6.1.2 Therapeutic HPV Vaccine

Cervical cancer is the second leading cause of death afflicting women worldwide; 40% of cervical patients develop persistent, recurrent, or widely metastatic disease. While a preventive vaccine now exists for HPV, there is a need for a therapeutic vaccine to treat existing cases of HPV, especially in resource-poor areas where access to the preventive vaccine is limited.

Cellular immune responses are believed to be critical for effective immune response to cancer; accordingly, EpiVax is pursuing the development of an 'immunotherapeutic vaccine' which would focus on the proteins primarily expressed either before or during carcinogenesis (E1 and E2, E6, and E7). In order to maximize immunogenicity across HPV subtypes, our strategy involves analyzing variant strains of HPV protein sequences, using (1) EpiMatrix to identify Class I and Class II HLA motif matches, (2) Conservatrix to identify those motif matches that are conserved, and (3) Epi-Assembler to weave together the conserved immunogenic sequences into a full immunogenic consensus sequence (ICS) antigen. A full ICS vaccine antigen would retain the fundamental structure of its naturally occurring counterparts; however, it will contain more and better epitopes than would occur in any such one counterpart. The authors have identified five conserved epitopes in E1 and E2, which

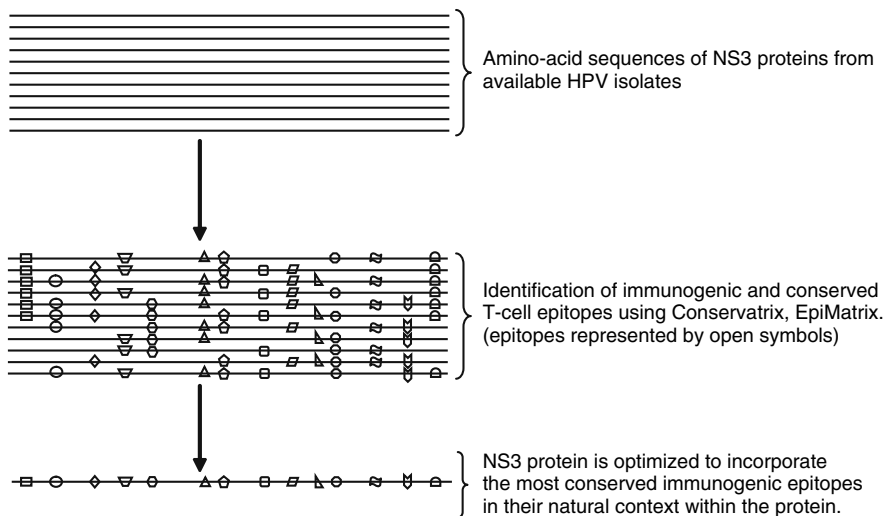


Fig. 5 Illustration of our novel strategy to generate an HPV vaccine candidate. The putative epitopes identified during this analysis will be combined to form several consensus sequences, which retain the fundamental structure of these HPV proteins but which also contain an unnaturally large number of conserved T cell epitopes

have stimulated significant responses in ELISpot IFN-g assays. In the proposed vaccine, these epitopes and others will be incorporated in their natural context within the proteins, which could be delivered as DNA, proteins, or a prime-boost combination. By preserving the natural flanking regions surrounding our epitopes, we hope to retain, in large part, the natural processes surrounding HPV protein degradation, transport, and presentation as they occur during natural infections. The ICS approach described here is the same as that illustrated in Fig. 5, but extending over the full natural length of the protein. For a more in-depth review of a similar approach being pursued for influenza, see McMurry et al. (2008).

The EpiVax HPV vaccine illustrates yet another aspect of vaccine design: ‘megatope’ proteins, re-engineered to increase the epitope content. This approach already had some success (Okazaki et al. 2006).

7 Advantages and Disadvantages of IDV

In the case of variable viruses such as HCV, influenza, and HIV, one limitation of conventional vaccination, and of natural infection, is that the immune system often focuses strongly on the most mutable immunogens. IDVs can be constructed from alternative antigens, which are more conserved or more protective, circumventing this problem (Russell and Liew 1979; Scherle and Gerhard 1986; Scherle and Gerhard 1988; Santra et al. 2002; Subbramanian et al. 2003).

In addition, broadening the T cell repertoire might make it possible to impair viral escape and decrease viral loads sufficiently to disrupt transmission.

Epitope-driven vaccines also offer distinct advantages over vaccines encoding whole protein antigens, since epitopes are safe and can be packaged into relatively small delivery vehicles. The epitope-driven approach offers platform independence: a delivery vehicle (peptide, DNA, multi-epitope construct) can be modified or selected midway into the development process. Multiple conserved epitopes, in addition to augmenting the efficacy of a preventive vaccine, could provide a broad and universal cellular immunity, known to be crucial for containment of infection, although perhaps ineffective for protection against infection.

Despite these advantages, there are a number of reasons that a given pathogen-directed, epitope-based vaccine might fail to reach clinical trials or protect humans: (1) the limited number of epitopes expressed by the vaccine (i.e., poor payload quantity); (2) limited conservation of epitopes (leading to limited coverage of variant clinical isolates) (3) the limited HLA population coverage (i.e., poor payload quality); (4) suboptimal vaccine delivery; and/or (5) the dearth of suitable animal models.

In addition, the concept of epitope-driven vaccines is relatively novel. Complete genome sequences have been available for only a little more than a decade now and the tools to process the data for vaccine design are only newer. Experimental validation needed to push forward these vaccines into clinical trials is now emerging and promises to enable epitope-based vaccines to claim a prominent place in the vaccine world.

8 Future Research

The technologies needed to identify immunostimulatory antigens and epitopes from pathogen genomes are already well developed. The principle focus of future research in this area will likely be in fine-tuning these technologies and expanding them to tailor immune responses in individuals. For example, development of epitope mapping algorithms for DQ and DP Class II HLA alleles will make it possible to completely characterize immunomes. This information will make it possible to generate comprehensive individual T cell epitope measures (iTEM) based on an individual's HLA genetic make-up and allow researchers to identify a priori clinically important epitopes and screen clinical cohorts for subjects that are more likely to develop targeted immune responses.

Furthermore, genome-mapping tools that are currently available are not yet useful for discovering B cell epitopes, whether from proteins or from non-protein components such as carbohydrates or lipid antigens. Immunoinformatics tools that are currently available cannot be used to accurately predict conformational (B cell) epitopes that interact with antibody, although such tools are being refined (Enshell-Seijffers et al. 2003). Thus, the immunogens

identified using *in silico* approaches must be evaluated *in vitro* and also in appropriate challenge models, prior to progressing to vaccine trials.

Protective immune response probably also involves some engagement of the innate immune system; it has been impossible to differentiate between effective and non-protective epitopes. Cytokine milieu may affect the outcome of immunization; thus, a limited number of toll-receptor agonists (Imler and Hoffmann 2001) have been identified and these are under study in conjunction with IDVs. In the future, toll-receptor signaling ‘pathogen-associated molecular patterns’ (PAMPS) might also be modeled and selected using immunoinformatics tools.

Besides antigen identification, the success of IDVs relies heavily on delivery technologies. These areas continue to independently mature and provide important lessons to epitope-based vaccine design. The major areas of research to watch include biological macromolecule (including cytokines), lipopeptide, and polysaccharide adjuvants and particulate (liposomes, exosomes, virosomes, nanoparticles) and cell-based delivery systems.

9 Conclusion

The development of safe and effective vaccines against emerging infectious diseases such as influenza, both seasonal and pandemic, HIV, and TB, in addition to cancers associated with infectious pathogens such as HBV and HCV, is an urgent and achievable public health priority. In addition, vaccines for the prevention and treatment of cancer hold enormous promise for human health.

The threat of bioterrorism following the events of September 11, 2001, provided vaccinologists with a persuasive argument for more rapid development of vaccines against viral and bacterial pathogens that are now included on the NIH Category A-C Biopathogen list (<http://www3.niaid.nih.gov/topics/BiodefenseRelated/Biodefense/research/CatA.htm>). ‘Emerging infectious diseases’ were added to the vaccine wish list following the outbreak of Severe Acute Respiratory Syndrome (SARS) in Guangdong China in 2002. Indeed, only a few months following the publication of the SARS-Coronavirus (SARS-CoV) genome (Marra et al. 2003; Rota et al. 2003), researchers began to map vaccine components using new bioinformatics and immunoinformatics tools, coupled with improved immunology techniques and specialized animal models. New vaccines based on this approach are currently being evaluated in animal models, less than a year from the start of the epidemic.

Future vaccine approaches may need to move away from ‘whole’ protein vaccines for a wide range of reasons. Multiple antigen or epitope vaccinations such as the approach illustrated here could be one way to elicit the sort of strong TH1 response necessary to pathogens following infection, in the context of a therapeutic vaccine. This approach could also be useful for a wide range of pathogens for which genomes have been partially or completely mapped. As described in this chapter, our group is actively pursuing the development of

epitope-driven vaccines for HIV (De Groot et al. 2005; Koita et al. 2006), *Francisella tularensis*, *Helicobacter pylori*, and smallpox. We have progressed from genome-derived epitope mapping to challenge studies in less than one year for some of these vaccine development programs.

Epitope-based and whole antigen IDVs are now just beginning to enter clinical trials, but this relative disadvantage may be cured with the tincture of time. One reason for the relative paucity of IDVs in clinical development is that the immunoinformatics tools for developing these vaccines have really only evolved in the last 10 to 15 years. The average length of time to develop a vaccine may be 20 years or more. While immunoinformatics tools are useful for accelerating the discovery and pre-clinical stage of vaccine development, testing vaccines in animal models and developing clinical trials is a lengthy process. It is likely that IDV and epitope-based IDV will begin to enter clinical trials and emerge on the market in greater numbers in 5 to 10 years.

References

- Ahlers, J. D., Belyakov, I. M., Thomas, E. K., and Berzofsky, J. A. 2001. High-affinity T helper epitope induces complementary helper and APC polarization, increased CTL, and protection against viral infection. *J Clin Invest.* 108:1677–1685.
- An, L. L., Rodriguez, F., Harkins, S., Zhang, J., and Whitton, J. L. 2000. Quantitative and qualitative analyses of the immune responses induced by a multivalent minigene DNA vaccine. *Vaccine.* 18:2132–2141.
- An, L. L., Whitton, J. L. 1997. A multivalent minigene vaccine, containing B-cell, cytotoxic T-lymphocyte, and Th epitopes from several microbes, induces appropriate responses in vivo and confers protection against more than one pathogen. *J Virol.* 71:2292–2302.
- Asjö, B., Stavang, H., Sørensen, B., Baksaas, I., Nyhus, J., and Langeland, N. 2002. Phase I trial of a therapeutic HIV type 1 vaccine, Vacc-4x, in HIV type 1-infected individuals with or without antiretroviral therapy. *AIDS Res Hum Retroviruses.* 18:1357–1365.
- Barnea, E., Beer, I., Patoka, R., Ziv, T., Kessler, O., Tzehoval, E., Eisenbach, L., Zavazava, N., and Admon, A. 2002. Analysis of endogenous peptides bound by soluble MHC class I molecules: a novel approach for identifying tumor-specific antigens. *Eur J Immunol.* 32:213–222.
- Bergmann, C. C., Tong, L., Cua, R., Sensintaffar, J., and Stohlgren, S. 1994. Differential effects of flanking residues on presentation of epitopes from chimeric peptides. *J Virol.* 68:5306–5310.
- Betts, M. R., Ambrozak, D. R., Douek, D. C., Bonhoeffer, S., Brenchley, J.M., Casazza, J.P., Koup, R. A., and Picker, L. J. 2001. Analysis of total human immunodeficiency virus (HIV)-specific CD4(+) and CD8(+) T cell responses: relationship to viral load in untreated HIV infection. *J Virol.* 75:11983–11991.
- Bond, K. B., Sriwanthana, B., Hodge, T. W., De Groot, A. S., Mastro, T. D., Young, N. L., Promadej, N., Altman, J. D., Limpakarnjanarat, K., and McNicholl, J. M. 2001. An HLA-directed molecular and bioinformatics approach identifies new HLA-A11 HIV-1 subtype E cytotoxic T lymphocyte epitopes in HIV-1-infected Thais. *AIDS Res Hum Retroviruses.* 20:703–717.
- Charo, J., Sundback M, Geluk, A., Ottenhoff, T., and Kiessling, R.. 2001. DNA immunization of HLA transgenic mice with a plasmid expressing mycobacterial heat shock protein 65 results in HLA class I- and II-restricted T cell responses that can be augmented by cytokines. *Hum Gene Ther.* 12:1797–1804.

- Chen, W., Norbury, C. C., Cho, Y., Yewdell, J. W., and Bennink, J. R. 2001. Immunoproteasomes shape immunodominance hierarchies of antiviral CD8(+) T cells at the levels of T cell repertoire and presentation of viral antigens. *J Exp Med.* 193:1319–1326.
- Crowe, S. R., Miller, S. C., Brown, D. M., Adams, P. S., Dutton, R. W., Harmsen, A. G., Lund, F. E., Randall, T. D., Swain, S.L., and Woodland, D. L. 2006. Uneven distribution of MHC class II epitopes within the influenza virus. *Vaccine.* 24:457–467.
- De Groot, A. S. 2006. Immunomics: Discovering New Targets for Vaccine and Therapeutics. *Drug Discov Today.* 11:203–209.
- De Groot, A. S. and Berzofsky, J. A. 2004. From Genome to Vaccine – New Immunoinformatics tools for vaccine design. *Methods.* 34:425–428.
- De Groot, A. S., and Martin, W. 2003. From immunome to vaccine: epitope mapping and vaccine design tools. *Novartis Found Symp.* 254:57–72.
- De Groot, A. S., Bishop, E., Khan, B., Marcon, L., Franco, J., Mayer, K., Carpenter, C., and Martin, W. 2004. Engineering immunogenic consensus T helper epitopes for a cross-clade HIV vaccine. *Methods.* 34:476–487.
- De Groot, A. S., Bosma, A., Chinai, N., Frost, J., Jesdale, B. M., Gonzalez, B. M., Martin, W., and Saint-Aubin, C. 2001. From genome to vaccine: in silico predictions, ex vivo verification. *Vaccine.* 19:4385–4395.
- De Groot, A. S., Clerici, M., Hosmalin, C. M., Hughes, A., Barnd, D., Hendrix, C. W., Houghten, R., Shearer, G. M. and Berzofsky, J. A. 1991. Human Immunodeficiency virus reverse transcriptase T helper epitopes identified in mice and humans: correlation with a cytotoxic T cell epitope. *J Infect Dis.* 164:1058–1065.
- De Groot, A. S., Jesdale, B. M., Szu, E., Schafer, J. R., Chicz, R. M., and Deocampo, G. 1997. An interactive Web site providing major histocompatibility ligand predictions: application to HIV research. *AIDS Res Hum Retroviruses.* 13:529–531.
- De Groot, A. S., Marcon, L., Bishop, E.A., Rivera, D., Kutzler, M., Weiner, D. B., and Martin, W. 2005. HIV vaccine development by computer assisted design: the GAIA vaccine. *Vaccine.* 23:2136–2148.
- De Groot, A. S., Sbai, H., Aubin, C. S., McMurry, J., and Martin, W. 2002. Immunoinformatics: Mining genomes for vaccine components. *Immunol Cell Biol.* 80:225–269.
- Depla, E., Van der Aa, A., Livingston, B. D., Crimi, C., Allosery, K., De Brabandere, V., Krakover, J., Murthy, S., Huang, M., Power, S., Babé, L., Dahlberg, C., McKinney, D., Sette, A., Southwood, S., Philip, R., Newman, M. J., and Meheus, L. 2008. Rational design of a multiepitope vaccine encoding T-lymphocyte epitopes for treatment of chronic hepatitis B virus infections. *J Virol.* 82:435–50.
- Dong, Y., Demaria, S., Sun, X., Santori, F. R., Jesdale, B. M., De Groot, A. S., Rom, W. N., and Bushkin, Y. 2004. HLA-A2-restricted CD8+ -cytotoxic- T cell responses to novel epitopes in *Mycobacterium tuberculosis* superoxide dismutase, alanine dehydrogenase, and glutamine synthetase. *Infect Immun.* 72: 2412–2415.
- Doytchinova, I. A., Taylor, P., and Flower, D. R. 2003. Proteomics in Vaccinology and Immunobiology: An Informatics Perspective of the Immunome. *J Biomed Biotechnol.* 2003:267–290.
- Elliott, S. L., Suhrbier, A., Miles, J. J., Lawrence, G., Pye, S. J., Le, T. T., Rosenstengel, A., Nguyen, T., Allworth, A., Burrows, S. R., Cox, J., Pye, D., Moss, D. J., and Bharadwaj M. A. 2008. Phase I trial of a CD8+ T Cell Peptide Epitope-based Vaccine for Infectious Mononucleosis. *J Virol.* 82:1448–1457.
- Enshell-Seiffers, D., Denisov, D., Groisman, B., Smelyanski, L., Meyuhas, R., Gross, G., Denisova, G., and Gershoni, J. M. 2003. The mapping and reconstitution of a conformational discontinuous B-cell epitope of HIV-1. *J Mol Biol.* 334:87–101.
- Falk, K., Rotzschke, O., Stevanovic, S., Jung, J., and Rammensee, H. G. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature.* 351:290–296.

- Frahm, N., Korber, B. T., Adams, C. M., Szinger, J. J., Draenert, R., Addo, M. M., Feeney, M. E., Yusim, K., Sango, K., Brown, N. V., SenGupta, D., Piechocka-Trocha, A., Simonis, T., Marincola, F. M., Wurcel, A. G., Stone, D. R., Russell, C. J., Adolf, P., Cohen, D., Roach, T., StJohn, A., Khatri, A., Davis, K., Mullins, J., Goulder, P. J., Walker, B. D., and Brander, C. 2004. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *J Virol.* 78:2187–2200.
- Franke, E. D., Sette, A., Sacci, J. Jr., Southwood, S., Corradin, G., and Hoffman, S. L. 2000. A subdominant CD8(+) cytotoxic T lymphocyte (CTL) epitope from the *Plasmodium yoelii* circumsporozoite protein induces CTLs that eliminate infected hepatocytes from culture. *Infect Immun.* 68:3403–3411.
- Fraser, C. K., Diener, K. R., Brown, M. P., and Hayball, J. D. 2007. Improving vaccines by incorporating immunological adjuvants. *Expert Rev Vaccines.* 6:559–578.
- Gahery, H., Daniel, N., Charmeteau, B., Ourth, L., Jackson, A., Andrieu, M., Choppin, J., Salmon, D., Pialoux, G., and Guillet, J. G. 2006. New CD4+ and CD8+ T cell responses induced in chronically HIV type-1-infected patients after immunizations with an HIV type 1 lipopeptide vaccine. *AIDS Res Hum Retroviruses.* 22:684–94.
- Ghosh, S., Pal, S., Das, S., Dasgupta, S. K., and Majumdar, S. 1998. Lipoarabinomannan induced cytotoxic effects in human mononuclear cells. *FEMS Immunol Med Microbiol.* 21:181–188.
- Gianfrani, C., Oseroff, C., Sidney, J., Chesnut, R. W., and Sette, A. 2000. Human memory CTL response specific for influenza A virus is broad and multispecific. *Hum Immunol.* 61:438–452.
- Godkin, A. J., Smith, K. J., Willis, A., Tejada-Simon, M. V., Zhang, J., Elliott, T., and Hill, A. V. 2001. Naturally processed HLA class II peptides reveal highly conserved immunogenic flanking region sequence preferences that reflect antigen processing rather than peptide-MHC interactions. *J Immunol.* 166:6720–6727.
- Graham, C. M., Barnett, B. C., Hartlmayr, I., Burt, D. S., Faulkes, R., Skehel, J. J., and Thomas, D. B. 1989. The structural requirements for class II (I-Ad)-restricted T cell recognition of influenza hemagglutinin: B cell epitopes define T cell epitopes. *Eur J Immunol.* 19:523.
- Gulukota, K. and DeLisi, C. 1996. HLA allele selection for designing peptide vaccines. *Genet Anal.* 13:81–86.
- Hernychova, L., Stulik, J., Halada, P., Macela, A., Kroca, M., Johansson, T., and Malina, M. 2001. Construction of a *Francisella tularensis* two-dimensional electrophoresis protein database. *Proteomics.* 1:508–515.
- Hozhutter, H. G., Frommel, C., and Kloetzel, P. M. 1999. A theoretical approach towards the identification of cleavage determining amino acid motifs of the 20S proteasome. *J. Mol Biol.* 286:1251.
- Imler, J. L., and Hoffmann, J. A. 2001. Toll receptors in innate immunity. *Trends Cell Biol.* 11:304–311.
- Ishioka, G. Y., Fikes, J., Hermanson, G., Livingston, B., Crimi, C., Qin, M., del Guercio, M. F., Oseroff, C., Dahlberg, C., Alexander, J., Chesnut, R. W., and Sette, A.. 1999. Utilization of MHC class I transgenic mice for development of minigene DNA vaccines encoding multiple HLA-restricted CTL epitopes. *J. Immunol.* 162:3915–3925.
- Itoh K. and Yamada, A. 2006. Personalized peptide vaccines: a new therapeutic modality for cancer. *Cancer Sci.* 97:970–976.
- Iversen, A. K., Stewart-Jones, G., Learn, G. H., Christie, N., Sylvester-Hviid, C., Armitage, A. E., Kaul, R., Beattie, T., Lee, J. K., Li, Y., Chotiarnwong, P., Dong, T., Xu, X., Luscher, M. A., MacDonald, K., Ullum, H., Klarlund-Pedersen, B., Skinhoj, P., Fugger, L., Buus, S., Mullins, J. I., Jones, E. Y., van der Merwe, P. A., and McMichael, A. J. 2006. Conflicting selective forces affect T cell receptor contacts in an immunodominant human immunodeficiency virus epitope. *Nat Immunol.* 7:179–189.
- Jameson, J., Cruz, J., and Ennis, F.A. 1998. Human cytotoxic T-lymphocyte repertoire to influenza A viruses. *J Virol.* 72:8682–8689.

- Johansson, B. E., Moran, T. M., and Kilbourne, E. D. 1987. Antigen-presenting B cells and helper T cells cooperatively mediate intravirionic antigenic competition between influenza A virus surface glycoproteins. *Proc Natl Acad Sci U S A.* 84:6869–6873.
- Kaufmann, S. H., Gulle, H., Daugelat, S., and Schoel, B. 1992. Tuberculosis and leprosy: attempts to identify T-cell antigens of potential value for vaccine design. *Scand J Immunol Suppl.* 11:85–90.
- Klitgaard, J. L., Coljee, V. W., Andersen, P. S., Rasmussen, L. K., Nielsen, L. S., Haurum, J. S., and Bregenholt S. 2006. Reduced susceptibility of recombinant polyclonal antibodies to inhibitory anti-variable domain antibody responses. *J Immunol.* 177:3782–3790.
- Koibuchi, T., Allen, T. M., Lichterfeld, M., Mui, S. K., O'Sullivan, K. M., Trocha, A., Kalams, S. A., Johnson, R.P., and Walker, B. D. 2005. Limited sequence evolution within persistently targeted CD8 epitopes in chronic human immunodeficiency virus type 1 infection. *J Virol.* 79:8171–8181.
- Koita, O. A., Dabitaou, D., Mahamadou, I., Tall, M., Dao, S., Toukara, A., Guiteye, H., Noumsi, C., Thiero, O., Kone, M., Rivera, D., McMurry, J. A., Martin, W., and De Groot, A. S. 2006. Confirmation of immunogenic consensus sequence HIV-1 T cell epitopes in Bamako, Mali and Providence, Rhode Island. *Hum Vaccin.* 2:119–128.
- Koren, E., De Groot, A. S., Jawa, V., Beck, K. D., Boone, T., Rivera, D., Li, L., Mytych, D., Koscec, M., Weeraratne, D., Swanson, S., and Martin W. 2007. Clinical validation of the “in silico” prediction of immunogenicity of a human recombinant therapeutic protein. *Clin Immunol.* 124:26–32.
- Kran, A. M., Sørensen, B., Nyhus, J., Sommerfelt, M. A., Baksaas, I., Bruun, J.N., and Kvale, D. 2004. HLA- and dose-dependent immunogenicity of a peptide-based HIV-1 immunotherapy candidate (Vacc-4x). *AIDS.* 18:1875–1883.
- Krieg, A. M., Yi, A. K., Schorr, J., and Davis, H. L. 1998. The role of CpG dinucleotides in DNA vaccines. *Trends Microbiol.* 6:23–27.
- Kuchroo, V. K., Das, M. P., Brown, J. A., Ranger, A. M., Zamvil, S. S., Sobel, R. A., Weiner, H. L., Nabavi, N., and Glimcher, L. H. 1995. B7-1 and B7-2 costimulatory molecule activate differentially the TH1/TH2 developmental pathways: application to autoimmune disease therapy. *Cell.* 80: 707–718.
- Kulkarni-Kale, U., Bhosle, S., and Kolaskar, A. S. 2005. CEP: a conformational epitope prediction server. *Nucleic Acids Res.* 33:W168–171.
- Larsson, P., Oyston, P. C., Chain, P., Chu, M. C., Duffield, M., Fuxelius, H. H., Garcia, E., Halltorp, G., Johansson, D., Isherwood, K. E., Karp, P. D., Larsson, E., Liu, Y., Michell, S., Prior, J., Prior, R., Malfatti, S., Sjostedt, A., Svensson, K., Thompson, N., Vergez, L., Wagg, J. K., Wren, B. W., Lindler, L. E., Andersson, S. G., Forsman, M., and Titball, R. 2005. W. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. *Nat Genet* 37:153–159.
- Lazarski, C. A., Chaves, F. A., Jenks, S. A., Wu, S., Richards, K. A., Weaver, J. M., and Sant, A. J. 2005. The kinetic stability of MHC class II peptide complexes is a key parameter that dictates immunodominance. *Immunity.* 23:29–40.
- Leen, A. M., Christin, A., Khalil, M., Weiss, H., Gee, A. P., Brenner, M. K., Heslop, H. E., Rooney, C. M., and Bollard, C. M. 2008. Identification of hexon-specific CD4 and CD8 T-cell epitopes for vaccine and monotherapy. *J Virol.* 82:546–554.
- Livingston, B. D., Newman, M., Crimi, C., McKinney, D., Chesnut, R., and Sette, A. 2001. Optimization of epitope processing enhances immunogenicity of multi-epitope DNA vaccines. *Vaccine.* 19: 4652–4660.
- Man, S., Newberg, M. H., Crotzer, V. L. Luckey, C. J., Williams, N. S., Chen, Y., Huczko, E. L. Ridge, J. P., and Engelhard V. H. 1995. Definition of a human T-cell epitope from influenza A non-structural protein 1 using HLA-A2.1 transgenic mice. *Int. Immunol.* 7:597–605.
- Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattri, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., Pandoh, P. K., Petrescu, A. S., Robertson, A. G., Schein, J. E., Siddiqui, A.,

- Smailus, D. E., Stott, J. M., Yang, G. S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T. F., Bowness, D., Czub, M., Drobot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G. A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R. C., Krajden, M., Petric, M., Skowronski, D. M., Upton, C., Roper, R. L. 2003. The Genome Sequence of the SARS-Associated Coronavirus. *Science* 300:1399–1404.
- Mathiassen, S., Lauemoller, S. L., Ruhwald, M., Claesson, M. H., and Buus, S. 2001. Tumor-associated antigens identified by mRNA expression profiling induce protective anti-tumor immunity. *Eur J Immunol.* 31:1239–1246.
- McMurry, J. A., Gregory, S. H., Moise, L., Rivera, D., Buus, S., and De Groot, A. S. 2007a. Diversity of Francisella tularensis Schu4 antigens recognized by T lymphocytes after natural infections in humans: Identification of candidate epitopes for inclusion in a rationally designed tularemia vaccine. *Vaccine.* 25:3179–3191.
- McMurry, J.A., Johansson, B.E., and De Groot, A.S. 2008. A call to cellular and humoral arms: enlisting cognate T cell help to develop broad-spectrum vaccines against influenza A. *Hum Vaccin.* 4(2):148–157.
- McMurry, J. A., Moise, L., Gregory, S. H., and De Groot, A. S. 2007b. Tularemia vaccines – an overview. *Med Health R I.* 90:311–314.
- McMurry, J., Sbali, H., Gennaro, M. L., Carter, E. J., Martin, W., and De Groot, A. S. 2005. Analyzing Mycobacterium tuberculosis proteomes for candidate vaccine epitopes. *Tuberculosis (Edinb).* 85:95–105.
- Menne, S., Maschke, J., Tolle, T. K., Lu, M., and Roggendorf, M. Characterization of T-cell response to woodchuck hepatitis virus core protein and protection of woodchucks from infection by immunization with peptides containing a T-cell epitope. 1997. *J Virol.* 71:65–74.
- Mo, X. Y., Cascio, P., Lemerise, K., Goldberg, A. L., and Rock, K. 1999. Distinct proteolytic processes generate the C and N termini of MHC class I-binding peptides. *J Immunol* 163:5851–5859.
- Moutaftsi, M., Peters, B., Pasquetto, V., Tschärke, D. C., Sidney, J., Bui, H. H., Grey, H., and Sette, A. 2006. A consensus epitope prediction approach identifies the breadth of murine T(CD8+) -cell responses to vaccinia virus. *Nat Biotechnol.* 24:817–819.
- Nara, P. L. and Lin, G. 2005. HIV-1: the confounding variables of virus neutralization. *Curr Drug Targets Infect Disord.* 5:157–170.
- Nardin, E. H., Oliveira, G. A., Calvo-Calle, J. M., Castro, Z. R., Nussenzweig, R. S., Schmeckpeper, B., Hall, B. F., Diggs, C., Bodison, S., and Edelman, R. 2000. Synthetic malaria peptide vaccine elicits high levels of antibodies in vaccines of defined HLA genotypes. *J Infect Dis.* 182:1486–1496.
- Nguyen, L., Chaowanachan, T., Vanichseni, S., McNicholl, J. M., Mock, P. A., Nelson, R., Hodge, T. W., van Griensven, F., Choopanya, K., Mastro, T. D., Tappero, J. W., and Hu, D. J. 2004. Frequent human leukocyte antigen class I alleles are associated with higher viral load among HIV type 1 seroconverters in Thailand. *J Acquir Immune Defic Syndr.* 37:1318–1323.
- Okazaki, T., Pendleton, C. D., Sarobe, P., Thomas, E. K., Iyengar, S., Harro, C., Schwartz, D., Berzofsky, J. A. 2006. Epitope enhancement of a CD4 HIV epitope toward the development of the next generation HIV vaccine. *J Immunol.* 176:3753–3759.
- Olsen, A. W., Hansen, P. R., Holm, A., and Andersen, P. 2000. Efficient protection against Mycobacterium tuberculosis by vaccination with a single subdominant epitope from the ESAT-6 antigen. *Eur J Immunol.* 30:1724–1732.
- Ostrowski, M., Galeota, J. A., Jar, A. M., Platt, K. B., Osorio, F. A., and Lopez, O. J. 2002. Identification of neutralizing and nonneutralizing epitopes in the porcine reproductive and respiratory syndrome virus GP5 ectodomain. *J Virol.* 76:4241–4250.
- Paina-Bordignon, P., Tan, A., Termijtelen, A., Demotz, S., Corradin G., and Lanzavecchia, A. 1989. Universally immunogenic T cell epitopes: promiscuous recognition by T cells. *Eur J Immunol.* 19:2237–2242.

- Pederson, T. 1999. The immunome. *Mol Immunol.* 36:1127–1128.
- Petrovsky, N., and Brusic, V. 2002. Computational immunology: The coming of age. *Immunol Cell Biology.* 80:248–254.
- Peyerl, F. W., Bazick, H. S., Newberg, M. H., Barouch, D. H., Sodroski, J., and Letvin, N. L. 2004. Fitness costs limit viral escape from cytotoxic T lymphocytes at a structurally constrained epitope. *J. Virol.* 78:13901–13910.
- Pietersz, G. A., Pouniotis, D.S., and Apostolopoulos, V. 2006. Design of peptide-based vaccines for cancer. *Curr Med Chem.* 13:1591–1607.
- Pizza, M., Scarlato, V., Masignani, V., Giuliani, M. M., Aricò, B., Comanducci, M., Jennings, G. T., Baldi, L., Bartolini, E., Capecchi, B., Galeotti, C. L., Luzzi, E., Manetti, R., Marchetti, E., Mora, M., Nuti, S., Ratti, G., Santini, L., Savino, S., Scarselli, M., Storni, E., Zuo, P., Broecker, M., Hundt, E., Knapp, B., Blair, E., Mason, T., Tettelin, H., Hood, D. W., Jeffries, A. C., Saunders, N. J., Granoff, D. M., Venter, J. C., Moxon, E. R., Grandi, G., and Rappuoli, R. 2000. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science.* 287:1816–1820.
- Plotnicky, H., Cyblat-Chanal, D., Aubry, J. P., Derouet, F., Klinguer-Hamour, C., Beck, A., Bonnefoy, J. Y., and Corvaia, N. 2003. The immunodominant influenza matrix T cell epitope recognized in human induces influenza protection in HLA-A2/K(b) transgenic mice. *Virology.* 309:320–329.
- Rajnavolgyi, E., Nagy, N., Thuresson, B., Dosztanyi, Z., Simon, A., Simon, I., Karr, R. W., Ernberg, I., Klein, E., and Falk, K.I. 1999. A repetitive sequence of Epstein-Barr virus nuclear antigen 6 comprises overlapping T cell epitopes which induce HLA-DR restricted CD4+ T lymphocytes. *Int Immunol.* 12:281–293.
- Rappuoli, R. and Covacci, A. 2003. Reverse vaccinology and genomics. *Science.* 302:602.
- Rodriguez, F., Whitton, J. L. 2000. Enhancing DNA immunization. *Virology.* 268:233–238.
- Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Peñaranda, S., Bankamp, B., Maher, K., Chen, M. H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J. L., Chen, Q., Wang, D., Erdman, D. D., Peret, T. C., Burns, C., Ksiazek, T. G., Rollin, P. E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Günther, S., Osterhaus, A. D., Drosten, C., Pallansch, M. A., Anderson, L. J., Bellini W. J. 2003. Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *Science.* 300:1394–1399.
- Rotzschke, O., Falk, K., Stevanovic, S., Jung, J., Walden, P., and Rammensee, H. G.. 1991 Exact prediction of natural T cell epitope. *Eur J Immunol.* 21:2891–2894.
- Russell, S. M. and Liew, F. Y. 1979. T cells primed by influenza virion internal components can cooperate in the antibody response to haemagglutinin. *Nature.* 280:147–148.
- Santra, S., Barouch, D. H., Kuroda, M. J., Schmitz, J. E., Krivulka, G. R., Beaudry, K., Lord, C. I., Lifton, M. A., Wyatt, L. S., Moss, B., Hirsch, V. M., and Letvin, N. L. 2002. Prior vaccination increases the epitopic breadth of the cytotoxic T-lymphocyte response that evolves in rhesus monkeys following a simian-human immunodeficiency virus infection. *J Virol.* 76:6376–6381.
- Schadeck, E. B., Partidos, C. D., Fooks, A. R., Obeid, O. E., Wilkinson, G. W., Stephenson, J. R., and Steward, M. W. 1999. CTL epitopes identified with a defective recombinant adenovirus expressing measles virus nucleoprotein and evaluation of their protective capacity in mice. *Virus Res.* 65:75–86.
- Scherle P. A. and Gerhard, W. 1986. Functional analysis of influenza-specific helper T cell clones in vivo. T cells specific for internal viral proteins provide cognate help for B cell responses to hemagglutinin. *J Exp Med.* 164:1114–1128.
- Scherle, P. A. and Gerhard, W. 1988. Differential ability of B cells specific for external vs. internal influenza virus proteins to respond to help from influenza virus-specific T-cell clones in vivo. *Proc Natl Acad Sci U S A.* 85:4446–4450.

- Schreiber, A., Humbert, M., Benz, A., and Dietrich, U. 2005. 3D-Epitope-Explorer (3DEX): localization of conformational epitopes within three-dimensional structures of proteins. *J Comput Chem.* 26:879–887.
- Sepkowitz, K. A. 2001. Tuberculosis control in the 21st century. *Emerg Infect Dis.* 7:259–262.
- Sette, A. and Sidney, J. 1998. HLA supertypes and supermotifs: a functional perspective on HLA polymorphism. *Curr Opin Immunol.* 10:478–482.
- Shahinian, A., Pfeffer, K., Lee, K. P., Kundig, T. M., Kishihara, T. M., Kishihara, A., Wakeham, A., Kawai, K., Ohashi, P. S., Thompson, C. B. and Mak, T. W. 1993. Differential T cell costimulatory requirements in CD28-deficient mice. *Science.* 261:609–612.
- Shastri, N., Serwold, T., and Gonzalez, F. 1995. Presentation of endogenous peptide/MHC class I complexes is profoundly influenced by specific C-terminal flanking residues. *J Immunol.* 155:4339–4346.
- Shirai, M., Arichi, T., Nishioka, M., Nomura, T., Ikeda, K., Kawanishi, K., Engelhard, V. H., Feinstone, S. M., and Berzofsky, J. A. 1995. CTL responses of HLA-A2.1-transgenic mice specific for hepatitis C viral peptides predict epitopes for CTL of humans carrying HLA-A2.1. *J Immunol.* 154:2733–2742.
- Sonnenberg, M. G., and Belisle, J. T. 1997. Definition of Mycobacterium tuberculosis culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. *Infect Immun.* 65:4515–4524.
- Southwood, S., Sidney, J., Kondo, A., del Guercio, M. F., Appella, E., Hoffman, S., Kubo, R. T., Chesnut, R. W., Grey, H. M., and Sette, A. 1998. Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol.* 160:3363–3373.
- Steere, A. C., Klitz, W., Drouin, E. E., Falk, B. A., Kwok, W. W., Nepom, G. T., and Baxter-Lowe, L. A. 2006. Antibiotic-refractory Lyme arthritis is associated with HLA-DR molecules that bind a Borrelia burgdorferi peptide. *J Exp Med.* 203:961–971.
- Subbramanian, R. A., Kuroda, M. J., Charini, W. A., Barouch, D. H., Costantino, C., Santra, S., Schmitz, J. E., Martin, K. L., Lifton, M. A., Gorgone, D. A., Shiver, J. W., and Letvin, N. L. 2003. Magnitude and diversity of cytotoxic-T-lymphocyte responses elicited by multiepitope DNA vaccination in rhesus monkeys. *J Virol.* 77:10113–10118.
- Thomson, S. A., Burrows, S. R., Misko, I. S., Moss, D. J., Coupar, B. E., and Khanna, R. 1998. Targeting a polyepitope protein incorporating multiple class II-restricted viral epitopes to the secretory/endocytic pathway facilitates immune recognition by CD4+ cytotoxic T lymphocytes: a novel approach to vaccine design. *J Virol.* 72:2246–2252.
- Thomson, S. A., Khanna, R., Gardner, J., Burrows, S. R., Coupar, B., Moss, D. J. and Suhrbier, A. 1995. Minimal epitopes expressed in recombinant polyepitope protein are processed and presented to CD8+ T cells: implications for vaccine design. *Proc Natl Acad Sci U S A.* 92:5845.
- Tobery, T. W., Dubey, S. A., Anderson, K., Freed, D. C., Cox, K. S., Lin, S., Prokop, M. T., Sykes K. J., Mogg, R., Mehrotra, D. V., Fu, T. M., Casimiro, D. R., and Shiver, J. W. 2006. A comparison of standard immunogenicity assays for monitoring HIV type 1 gag-specific T cell responses in Ad5 HIV Type 1 gag vaccinated human subjects. *AIDS Res HumRetroviruses.* 22:1081–1090.
- Toes, R. E., Nussbaum, A. K., Degermann, S., Schirle, M., Emmerich, N. P., Kraft, M., Laplace, C., Zwinderman, A., Dick, T. P., Muller, J., Schonfisch, B., Schmid, C., Fehling, H. J., Stevanovic, S., Rammensee, H. G. and Schild, H. 2001. Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J Exp Med.* 194:1–12.
- Tomlinson, A. J., Jameson, S., and Naylor, S. 1996. Strategy for isolating and sequencing biologically derived MHC class I peptides. *J Chromatogr A.* 744:273–278.
- Ueda, Y., Itoh, T., Nukaya, I., Kawashima, I., Okugawa, K., Yano, Y., Yamamoto, Y., Naitoh, K., Shimizu, K., Imura, K., Fuji, N., Fujiwara, H., Ochiai, T., Itoi, H., Sonoyama, T., Hagiwara, A., Takesako, K., Yamagishi, H. 2004. Dendritic cell-based immunotherapy of

- cancer with carcinoembryonic antigen-derived, HLA-A24-restricted CTL epitope: Clinical outcomes of 18 patients with metastatic gastrointestinal or lung adenocarcinomas. *int J Oncol.* 24:909–917.
- Valmori, D., Dutoit, V., Ayyoub, M., Rimoldi, D., Guillaume, P., Lienard, D., Lejeune, F., Cerottini, J. C., Romero, P., and Speiser, D. E. 2003. Simultaneous CD8⁺ T cell responses to multiple tumor antigen epitopes in a multi-peptide melanoma vaccine. *Cancer Immun.* 3:15.
- Van Kaer, L., Ashton-Rickardt, P. G., Eichelberger, M., Gaczynska, M., Nagashima, K., Rock, K. L., Goldberg, A. L., Doherty, P. C., and Tonegawa, S. 1994. Altered peptidase and viral-specific T cell response in LMP2 mutant mice. *Immunity* 1:533–541.
- Website for Category A-C pathogens: <http://www3.niaid.nih.gov/topics/BiodefenseRelated/Biodefense/research/CatA.htm> (last visited January 21, 2008).
- Wherry, E. J., Puorro, K. A., Porgador, A., and Eisenlohr, L. C. 1999. The induction of virus-specific CTL as a function of increasing epitope expression: responses rise steadily until excessively high levels of epitope are attained. *J Immunol.* 163:3735–3745.
- Willett, T. A., Meyer, A. L., Brown, E. L., and Huber, B. T. 2004. An effective second-generation outer surface protein A-derived Lyme vaccine that eliminates a potentially autoreactive T cell epitope. *Proc Natl Acad Sci U S A.* 101:1303–1308.
- York, I. A., Goldberg, A. L., Mo, X. Y., and Rock, K. L. 1999. Proteolysis and class I major histocompatibility complex antigen presentation. *Immunol Rev* 172:49–66.